Assimilation of Temperature and Wind Data in the Planetary Boundary Layer within GSI for Rapid Refresh

Dezso Devenyi[1,2], Tom Schlatter[1,2], Steve Weygandt[1], Stan Benjamin[1]

[1]Global Systems Division, Earth System Research Laboratory, NOAA
[2]Cooperative Institute for Research in Environmental Sciences, University of Colorado

December 15, 2006

1. Introduction.

The present Rapid Update Cycle (RUC) high frequency meteorological analysis and prediction system, run operationally at the National Centers for Environmental Prediction (NCEP), will soon be replaced by a new system known as the Rapid Refresh (RR). The RR will utilize the WRF model and the Gridpoint Statistical Interploation (GSI) analysis modules. The GSI was originally developed at NCEP for global model assimilation (known then as the Spectral Statistical Interpolation), but has been recently adapted for regional model use. It includes full satellite data assimilation capabilities, making it ideal for the domain expansion planned with the RR implementation. Use of the GSI for the RR application requires adding a number of specific features of the RUC data assimilation system to the GSI . These features include:

1. Cloud/hydrometeor analysis (blend background 3-d hydrometeor fields, GOES, METAR clouds and visibility)

2. 3-d assimilation of METAR temperature, dewpoint and wind using background PBL depth and land-use dependencies

3. Possible revision to balance, background errors, and scale-dependence

4. Quality Control (QC) buddy check, including RAOB/GPS precipitable water

5. Adjustment of soil temperature and moisture

This report summarizes progress on the 1st task, the assimilation of surface observations (especially of METAR data) in the planetary boundary layer (PBL) for RR. It includes discussion of two related aspects. Use of the anisotropic error covariance formulation in the GSI to assimilate surface observations throughout the PBL is discussed in Section 2. An assessment of the simple and advanced surface forward models (as revealed by the observation minus background statistics) for both the RUC and the GSI in included in Section 3. Present and future work is then briefly discussed in Section 4.

2. Use of anisotropic error covariance formulation

In the current RUC, pseudo-observations are generated within the PBL in order to spread the effect of surface data throughout the PBL but at the same time limit their effect above the PBL. The use of pseudo-observations provides a practical method for extending the surface observation impact throughout the PBL, while restricting the impact of surface observations above the PBL. A disadvantage is that artificial observations (with unknown error characteristics) are introduced into the analysis in an ad hoc manner, which can produce analysis features that deviate from the intended variational balance. Within the Gridpoint Statistical Interpolation (GSI) package, enhanced error covariance specification capabilities exists that allows for a more elegant approach to the problem of extending surface observation influence throughout the PBL. We have completed a set of experiments to evaluate the viability of this new approach, which use the model-predicted fields of virtual potential temperature to generate appropriate correlation lengths for wind and temperature fields.

a) Coding aspects.

The new approach for assimilating temperature and wind data throughout the PBL for the Rapid Refresh has been tested using the September 2006 version of GSI. Modification were made to the computer code from *module anisofilter*, which consists of the following subroutines:

6. init_anisofilter_reg
7. anprewgt_reg
8. get2berr_reg
9. get3berr_reg
10. read_background
11. get_background
12. isotropic scales
13. get_theta_corrl_lenghts
14. smther_one

Modifications to subroutines *get3berr_reg* and *get_theta_corrl_lenghts* (note typo in "lenghts") have been introduced to test for the most appropriate correlation lengths. The part of the code where scale tuning is introduced is seen below.

```
do k=1,nsig
   qlth_temp(k)=qlth_temp0
   qlth_wind(k)=qlth_wind0
   if (k.le.44) then
     qltv_temp(k)=qltv_temp0
     qltv_wind(k)=qltv_wind0
   else
     qltv_temp(k)=qltv_temp0*dthetabarz(k)/dthetabarz(44)*4._r_kind
     qltv_wind(k)=qltv_wind0*dthetabarz(k)/dthetabarz(44)*4._r_kind
   endif
  end do
```

Within this code section, qltv_temp and qlvtv_wind are already downweighted in the lowest 44 levels, which has the result of increasing the 3D anisotropy. Increasing the anisotropy causes greater adaptivity (stretching and compacting) of the influence region based on the relative spacing of the background isentropic surfaces in the various directions. Modifications to the above code for the PBL surface observation assimilation tests include introducing additional adaptivity in the lowest 5-10 levels. Subsequent tests will expand this to examine a broader parameter space.

Because subroutine *get_theta_corrl_lenghts* computes the correlation lengths for temperature (*qltv_temp(k)*) and wind (*qltv_wind(k)*) as a function of the model level and because it is easier to keep control on length scales in this subroutine than in subroutine *get3berr_reg*, the latter is used for control purposes only (*Manuel Pondeca, personal communication*). Smaller values of the correlation length imply a stronger anisotropy (adaptivity). Sharpening of the anisotropy should be done carefully to avoid any instability in the analysis problem.

b) Test results

Several tests have been performed over the CONUS domain using NAM prepbufr and satellite bufr data with a RUC 1-h forecast as background. Comparison of sample run times (wall clock) is seen below:

1) 'isotropic' run on GSD's supercomputer with 27 processors
STARTING DATE-TIME NOV 29,2006 18:30:47.967
ENDING DATE-TIME NOV 29,2006 18:39:21.670
ELAPSED TIME: ~8:34 min.

2) 'anisotropic' run on GSD's supercomputer with 19 processors
STARTING DATE-TIME NOV 29,2006 20:19:29.542
ENDING DATE-TIME NOV 29,2006 20:30:26.417
ELAPSED TIME: ~10:57 min.

For this relatively small (CONUS) domain, anisotropic normalization is not too demanding. Based on suggestions by Manual Pondeca of NCEP (personal communication), two factors for reducing the computation time for the anisotropic version were employed: 1) setting normal = -200 and 2) using a coarsening factor of 4 on the filter. Normalization may lead to a more significant increase in run time over a large domain.

In the following, some test results are presented both from single-observation studies and real cases. Single-observation experiments are valuable for comparing the spatial "footprint" of the observation influence region for different analysis formulations. For the present purpose, individual 1°C temperature observation increments are introduced near the surface and also at 850 hPa. The impact on temperature and wind fields is described below.

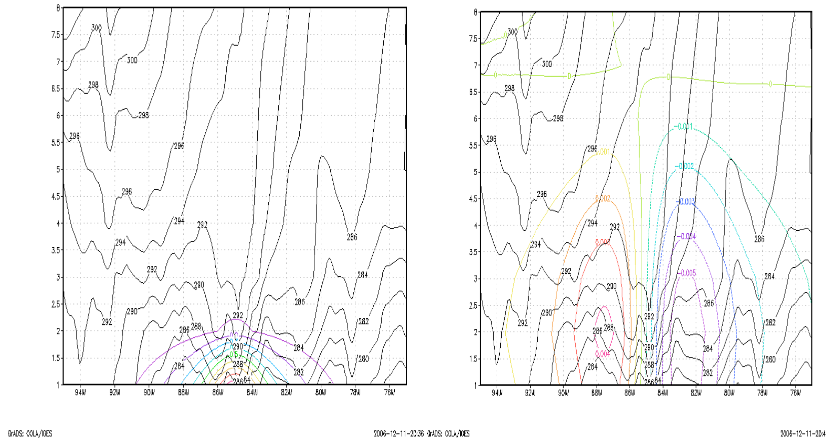

Fig. 1 Induced temperature (left) and u-component of wind (right) analysis increments at the lowest 8 model levels when a 1^{o}C increment in the temperature is applied close to the surface.. Anisotropy option is turned off.

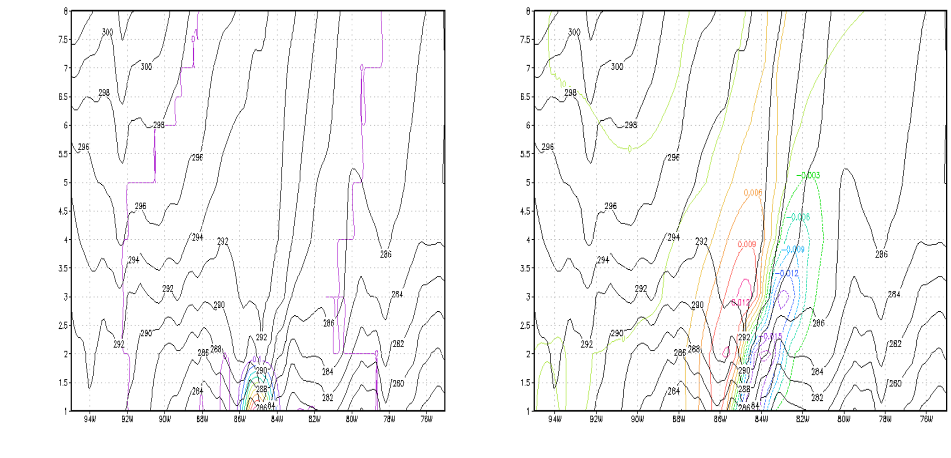

Fig. 2. Induced temperature (left) and u-component of wind (right) analysis increments at the lowest 8 model levels when a 1^{o}C temperature increment is applied close to the surface. Anisotropy option is used.

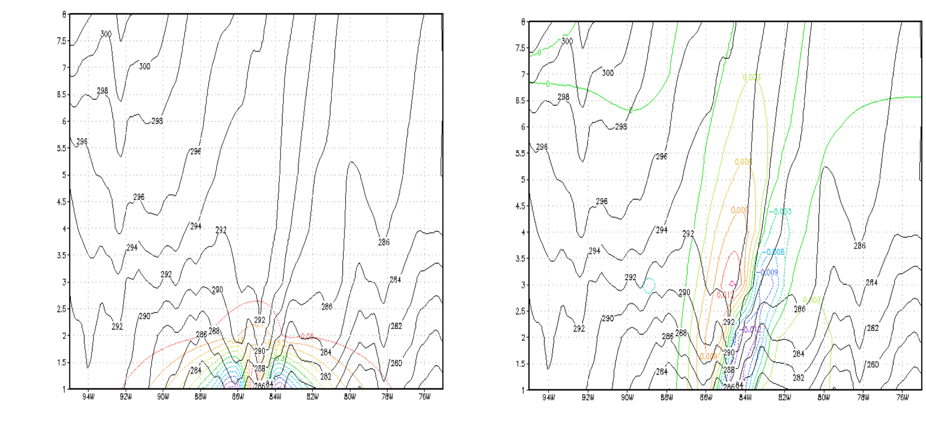

Fig. 2. Difference between isotropic and non-isotropic cases in induced temperature (left) and u-component of wind (right) analysis increments at the lowest 8 model levels when a 1^{o}C temperature increment is applied close to the surface.

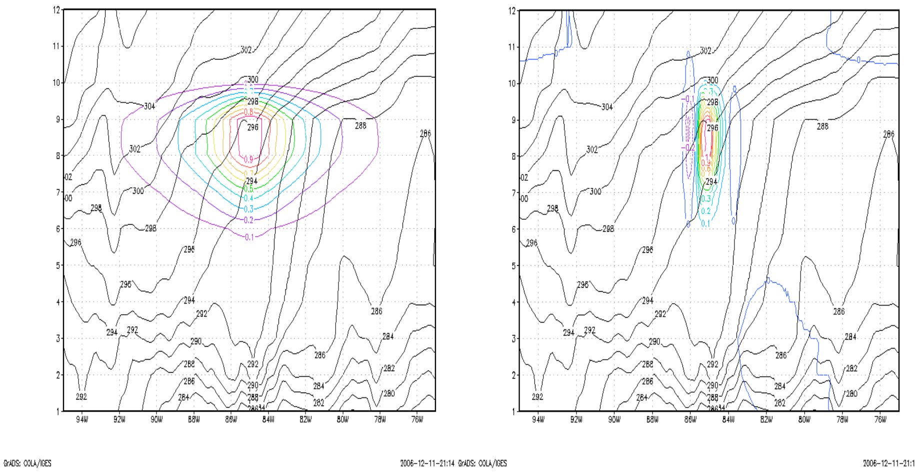

Fig. 3. Induced temperature with isotropy (left) and anisotropy (right) analysis increments at the lowest 12 model levels when a 1^{o}C temperature increment is applied at 850 hPa.

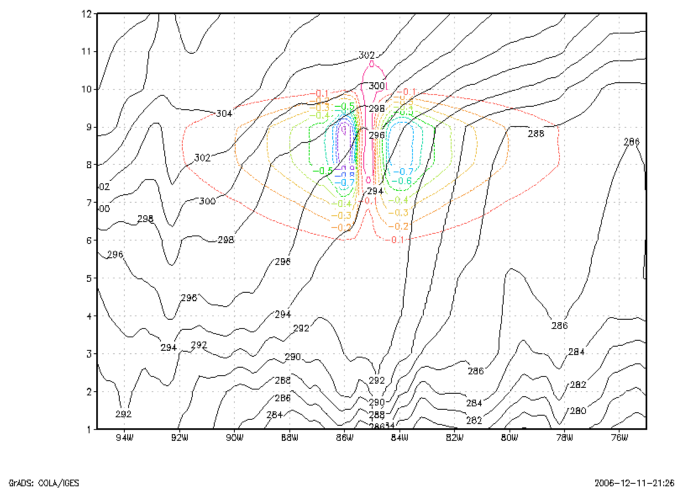

Fig. 4. Difference in induced temperature with isotropic and anisotropic analysis increments at the lowest 12 model levels when a 1^{o}C observation is applied at 850 hPa.

In all the previous anisotropic experiments an additional restriction was applied on the background scales using a linearly increasing weight profile from a value of 0.2 at level 1, increasing up to 1.0 at level 5. In the next case a similar weighting scheme was applied with the upper value of 1 achieved at level 10.

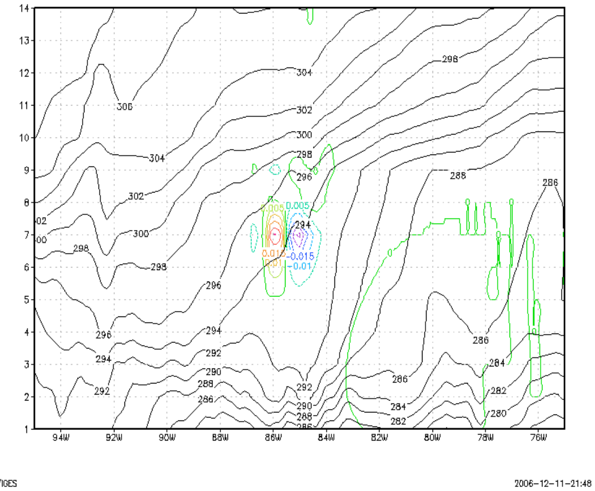

Fig. 5. Difference in induced temperature with differently weighted anisotropy analysis increments at the lowest 14 model levels when a 1°C observation increment is applied at 850 hPa (about model level 7).

Results from single-observation experiments presented above for the case of a single temperature observation, and similar results (not shown) for single wind observations prove the basic concept of application of correlation scales weighted according to the isentropic background. Numerically, GSI optimization performed well in all experiments, and the introduction of vertical weighting did not cause any computational instability. In the following, some real data experiments allow a comparison of isotropic and anisotropic background error cases. The first case is for 2006 April 11, 12 UTC, a case used at GSD for general GSI testing, including impact from radar radial winds, satellite radiance data, cloud drift wind, etc.

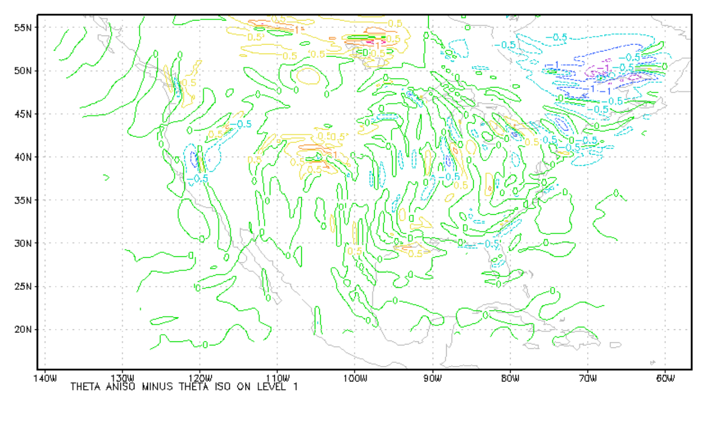

Fig . 6. Difference between analyzed anisotropic and isotropic temperature fields at model level 1 for case 2006 April 11, 12UTC. Contour interval is 0.5°C

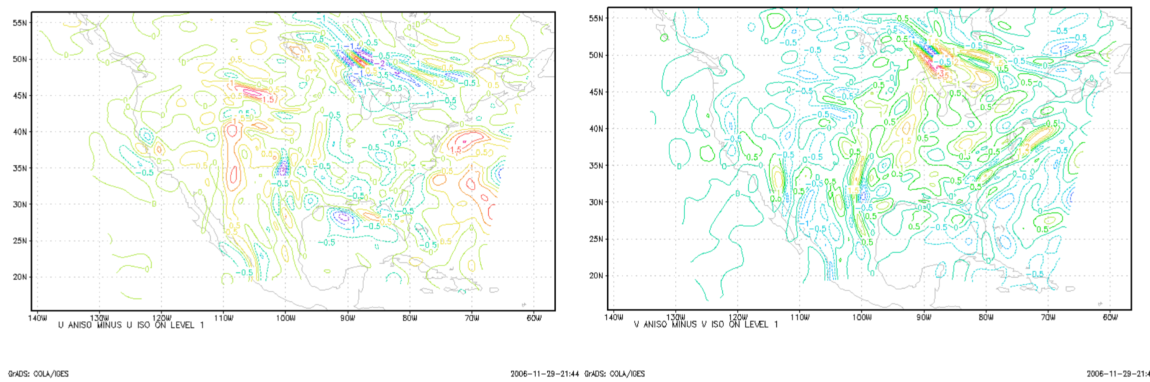

Fig . 7. Difference between analyzed anisotropic and isotropic u-components (left) and v-components of wind at model level 1 for case 2006 April 11, 12UTC. Contour interval is 0.5 m/s.

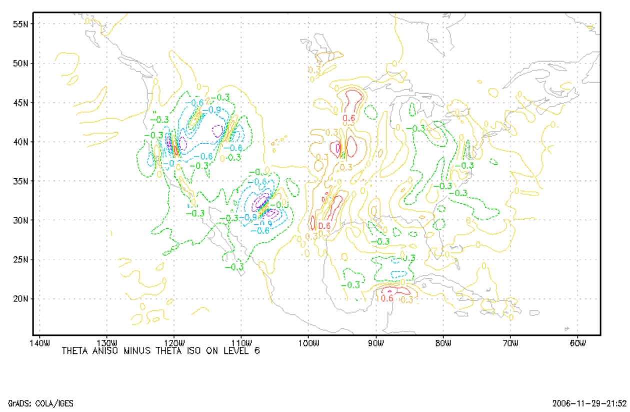

Fig . 8. Difference between analyzed anisotropic and isotropic temperature fields at model level 6 for case 2006 April 11, 12UTC. Contour interval is 0.3°C.

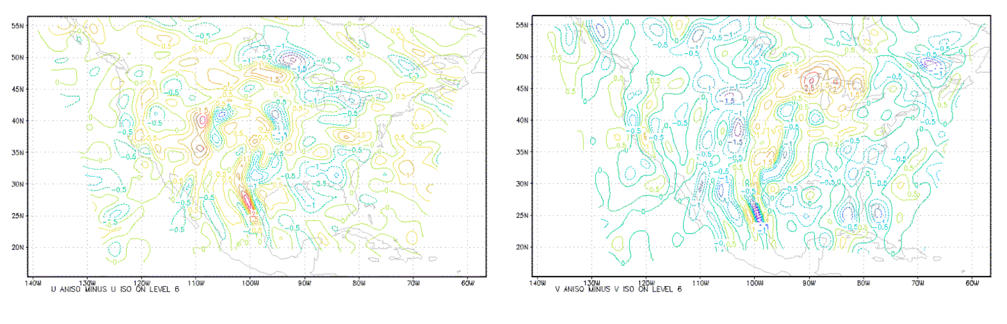

Fig . 9. Difference between analyzed anisotropic and isotropic u-components (left) and v-components of wind at model level 6 for case 2006 April 11, 12UTC. Contour interval is 0.5 m/s.

Similar results were realized for the other two cases investigated in detail: 2006 July 14, 12 UTC and 2006 July 15, 00 UTC. As an example, level 5 differences are presented below for 2006 July 15, 00 UTC.

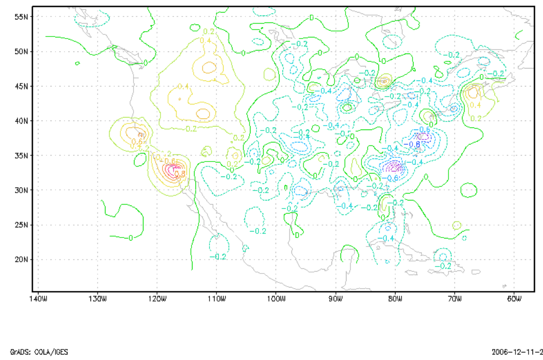

Fig . 10. Difference between analyzed anisotropic and isotropic temperature fields at model level 5 for case 2006 July 15, 00 UTC. Contour interval is 0.2°C.

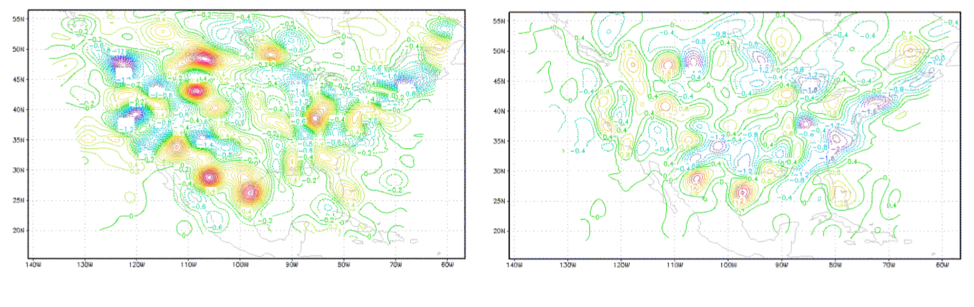

Fig . 11. Difference between analyzed anisotropic and isotropic u-components (left) and v-components of wind at model level 5 for case 2006 July 15, 00 UTC. Contour interval is 0.2 m/s for the u-component and 0.4 m/s for the v-component.

Results presented above in real data cases demonstrate the utility of the anisotropic approach in the analysis of lower level (PBL) fields.

c) Conclusions for anisotropic error covariance tests

Test results with single observations and full real data sets suggest the viability of correlation scale estimation based on forecasted isentropic fields. It helps to restrict observation impact to required analysis domain.
Further experiments are planned in the real time ARW cycle and for the extension of analysis domain to the one planned for RR.

3. Observation minus background differences

The starting point for every analysis is a short-range forecast, called a *background.* The analysis corrects the background based upon the available observations. The difference between observed values and the background is one measure of the accuracy of the background.

A key feature of the RUC 3DVAR analysis is the use of all available surface observations (provided they meet data quality control standards) to provide the most complete update of the near-surface background field. This is facilitated by the use of a similarity theory-based forward model (used to map the background field to the observations), in addition to the standard forward model (interpolation from the lowest model level). For the RUC analysis, the similarity forward model results in smaller observation minus background (O-B) differences for surface temperature observations (compared to use of the standard forward model), thereby allowing the surface temperature information to be used more effectively. Analogous similarity-theory-based and standard interpolation-based forward models exist within the GSI analysis package. An important part of the GSI testing and development work is the documentation of the O-B statistics for both forward models (standard and similarity) for the RUC and GSI analyses. These comparisons are presented below.

a) Comparison of standard vs. similarity forward models for RUC analysis

Modifications were made to the Rapid Update Cycle (RUC) surface analysis to make it more directly comparable to the Gridpoint Statistical Interpolation (GSI) surface analysis. O-B statistics were then stratified by pressure layer in order to take into account different station elevations, and also stratified by observation type, thereby recognizing that different instrument packages have different error characteristics. The three surface observation types initially considered were METARS, surface buoys, and surface mesonets.

The data assimilation algorithm matches each observed value with the corresponding value predicted by the model. In the case of the RUC, the prediction is a one-hour forecast. Because the model gridpoints do not lie at the observation locations, the model forecast has to be interpolated to the location of the observation. This is done in two ways: by straightforward interpolation from the lowest model level (just five meters above the model terrain) and in a more sophisticated way, taking into account similarity theory. Similarity theory tries to explain how temperature, moisture, and wind vary with height close to the ground. For example, if the wind observation is taken from a ten-meter tower (standard practice for METAR observations), similarity theory would say how the wind varied between ten meters and the closest model surface as a function of low-level stability.

This section of the report discusses a comparison between the two different methods, straight interpolation and similarity theory, for determining the background value at the location of the observation. We considered four variables: virtual temperature, pseudo relative humidity [obs-minus-background difference in mixing ratios normalized (divided by) the saturation mixing ratio], and the u- and v-components of the wind. In addition, we computed obs-minus-background differences for wind speed and the magnitude of the vector difference between the observed and background wind.

Here are sample results for METAR observations for a single analysis from last summer (0000 UTC 15 July 2006):

Virtual temperature [K]
Straight interpolation–black; similarity theory–blue

	>1000 mb	999-900 mb	899-800 mb	799-600 mb
Number of obs	501	1022	148	41
Mean	0.60	0.02	0.59	0.33
RMS	1.96	1.78	2.20	2.49
Mean	0.20	-0.52	-0.20	-0.33
RMS	1.72	1.70	2.06	2.58

Use of similarity theory brought the background closer to the observations than just straight interpolation. The rms differences in blue are significantly smaller. The differences grow with surface elevation (decreasing pressure).

Pseudo relative humidity (%)
Straight interpolation–black; similarity theory–blue

	>1000 mb	999-900 mb	899-800 mb	799-600 mb
Number of obs	499	1009	138	30
Mean	-3	-3	0	-1
RMS	8	8	5	6
Mean	-4	-4	0	-1
RMS	8	8	6	7

Use of similarity theory did not improve the scores for pseudo relative humidity.

u-component of the wind (m/s)
Straight interpolation–black; similarity theory–blue

	>1000 mb	999-900 mb	899-800 mb	799-600 mb
Number of obs	517	1018	145	41
Mean	-0.09	0.14	0.23	0.00
RMS	1.95	1.88	3.08	3.91
Mean	-0.15	0.07	0.19	-0.05
RMS	2.02	1.91	3.13	3.97

The obs-minus-background rms differences grow with station elevation (decreasing station pressure). Use of similarity theory leads to somewhat larger rms differences.

v-component of the wind (m/s)
Straight interpolation–black; similarity theory–blue

	>1000 mb	999-900 mb	899-800 mb	799-600 mb
Number of obs	517	1018	145	41
Mean	0.16	0.05	0.38	0.12
RMS	2.07	1.87	2.82	2.85
Mean	0.06	0.01	0.36	0.05
RMS	2.05	1.91	2.86	2.94

Except for stations near sea level, use of similarity theory leads to larger rms differences.

Except in the case of virtual temperature, similarity theory is worse than straight interpolation in estimating the observed value from the background. Part of the reason is that the similarity theory employed lacks theoretical rigor and contains some *ad hoc* assumptions. Even so, it is not yet clear that an improved similarity theory is the answer. One key feature in the RUC model and analysis system is the specification of model surfaces very close to the ground. This lessons the need for the more sophisticated similarity theory-based forward model.

b) Comparison of standard vs. similarity forward models for GSI analysis

Using a matched case (0000 UTC 15 July 2006, similar O-B statistics for both the standard and similarity-theory based forward models were computed for the GSI analysis. One factor that complicates a direct comparison between the RUC and GSI results is the vastly different number of observations. A factor in this difference is a difference in the way the RUC and GSI apply quality control (QC) to the observations. RUC uses a "buddy check" procedure, whereby observations that fail the QC are removed from the analysis. In contrast the GSI uses a variational down-weighting procedure, in which all observations are retained, but the influence of suspect observations is greatly reduced. Thus, for this paper, we will only draw qualitative conclusions about similarities and differences in the results for RUC vs. GSI. We are working to complete additional experiments in which we force an identical subset of observations to be used by both the RUC and GSI analyses.

Here are sample results for METAR observations for a single analysis from last summer:

Virtual temperature [K]
Straight interpolation–black; similarity theory–blue

	>1000 mb	999-900 mb	899-800 mb	799-600 mb
Number of obs	803	1927	219	73
Mean	-2.23	-2.29	0.51	0.60
RMS	3.07	3.10	2.71	3.35
Mean	-2.78	-3.21	-1.26	-2.23
RMS	3.49	3.76	3.16	4.36

In contrast to the RUC, use of the similarity theory yielded O-B statistics that were slightly larger than those obtained from straight interpolation. Also of note, a significant warm bias in the background field is noted in the mean differences. The cause of this bias is currently under investigation.

The GSI results for moisture and wind indicate no difference between the two forward model procedures, as the similarity-theory based forward model is inactive in the GSI for these fields. Qualitative comparison against the RUC indicates similar rms errors for moisture and winds. Note that wind differences are for vector quantities.

Pseudo relative humidity (%)

Straight interpolation–black; similarity theory–blue

	>1000 mb	999-950 mb	949-900 mb	899-850 mb
Number of obs	801	1675	246	134
Mean	-0.5	-0.4	0.1	0.3
RMS	7.4	7.0	6.3	5.5
Mean	-0.5	-0.4	0.1	-0.3
RMS	7.4	7.0	6.3	5.5

UV- wind (m/s)

Straight interpolation–black; similarity theory–blue

	>1000 mb	999-900 mb	899-800 mb	799-600 mb
Number of obs	715	1956	204	71
Mean	-1.13	-0.69	-0.09	-1.33
RMS	3.31	3.08	4.76	6.42
Mean	-1.13	-0.69	-0.09	-1.33
RMS	3.31	3.08	4.76	6.42

These results are preliminary, and further assessment is ongoing

5. Present and future work

Work continues to more fully evaluate the GSI surface analysis package run with the RUC-like modifications. The goal of this work is to verify that the GSI as formulated for the RR duplicates the desired RUC-specific aspects as described in section 1. Results to date suggest that the anisotropic capability within the GSI provides a viable alternative to the previous use of 'pseudo-obs' employed within the RUC analysis. Our preliminary assessment of O-B statistics for RUC and GSI reveal some differences that must be further examined. In particular the GSI may benefit from modifying the vertical grid structure to position the lowest model level closer to the surface. This change will be tested soon.

Additional work is ongoing for the implementation of two other RUC-specific features: 1) nudging of the soil temperature based on the surface temperature increment, and 2) modification of near-surface analysis increments in coastal regions based on the proximity to the coastline. Additional evaluation and testing is also ongoing using the real-time CONUS and full-domain cycles now running on ESRL computers.

5. Acknowledgements.

Special thanks to Drs. R. James Purser and Manuel Pondeca of NCEP for sharing ideas about anisotropic filtering and for their helpful suggestions and excellent cooperation.